Hindawi
Journal of Applied Mathematics
Volume 2020, Article ID 3108056, 9 pages
https://doi.org/10.1155/2020/3108056

Research Article

Theoretical Aspect of Diagonal Bregman Proximal Methods

S. Kabbadj

Department of Mathematics, Faculty of Sciences of Meknes, University Moulay Ismail, B.P. 11201, Meknes, Morocco

Correspondence should be addressed to S. Kabbadj; kabbajsaid63@yahoo.com

Received 4 August 2019; Revised 29 October 2019; Accepted 30 November 2019; Published 17 January 2020

Academic Editor: Quanke Pan

In this paper, we propose and study a diagonal inexact version of Bregman proximal methods, to solve convex optimization problems with and without constraints. The proposed method forms a unified framework for existing algorithms by providing others.

1. Introduction

Let f_i: $R^d \longrightarrow R$ $(i = 0, 1, \ldots, m)$ convex functions and C the nonempty subset of R^d are defined by

$$C = \left\{ x \in R^d : f_i(x) \le 0, i = 1, \ldots, m \right\}. \quad (1)$$

Let us consider the problem of convex optimization:

$$(P): \min \{ f_0(x), x \in C \}. \quad (2)$$

To solve (P), many authors [1–7] have combined the exterior penalty methods with the proximal method (PM) defined by

$$x^k \in \varepsilon_k - \text{Arg}\min \left\{ f(\cdot) + \frac{1}{2\lambda_k} \left\| \cdot - x^{k-1} \right\|^2 \right\}, \quad (3)$$

where $f \in \Gamma_0(R^d)$ is set of proper closed convex functions on R^d. PM and its variants have been studied by several authors [6, 8–13]. In this labor, we generalize this process by introducing Bregman's distance $D_h(.,.)$ defined by

$$D_h(x, y) := h(x) - h(y) - \langle x - y, \nabla h(y) \rangle, \quad (4)$$

where h is Bregman's function [14].

In order to solve (P), we study the coupling of the methods of the exterior penalty with the diagonal inexact version of the Bregman proximal methods defined by

$$x^k \in \varepsilon_k - \text{Arg}\min \left\{ f(\cdot) + \lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right) \right\}. \quad (5)$$

The exact version PMD is defined by

$$x^k = \arg\min \left\{ f(\cdot) + \lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right) \right\}, \quad (6)$$

has been studied by several authors [15–18].

We propose and study a diagonal inexact version of the Bregman proximal method, which we call DBPM, defined by

$$x^k \in \varepsilon_k - \text{Arg}\min \left\{ f^k(\cdot) + \lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right) \right\}, \quad (7)$$

where the sequence $\{f^k\}_k \subset \Gamma_0(R^d)$ is given and approaches f.

By introducing the penalty functions in DBPM, we deduce a solution of (P).

If $f^k = f \forall k$, the proposed method appears as an inexact version of (6) and solves the problem of convex optimization without constraints:

$$(P'): \min \left\{ f(x), x \in R^d \right\}. \quad (8)$$

For $h(\cdot) = (1/2)\|\cdot\|^2$, DBPM coincides with diagonal proximal method of Alart and Lemaire [1] as well as the penalization method given by Auslender [2].

2. Preliminary

In this section, we remind some theoretical properties of the approximations called entropic studied by Kabbadj in [17]. This study covers the properties of regularity and approximations of the Moreau–Yosida approximations [19]. These results are necessary for the analysis of the methods proposed in Section 3.

Let S be an convex open subset of R^d and $h\colon \overline{S} \longrightarrow R$. We define $D_h(.,.)$ by

$$\forall x \in \overline{S}, \forall y \in S\colon D_h(x,y) := h(x) - h(y) - \langle x-y, \nabla h(y)\rangle. \tag{9}$$

Let us consider the following hypotheses:

H_1: h is continuously differentiable on S.

H_2: h is continuous and strictly convex on $\overline{S}$.

H_3: $\forall r \geq 0, \forall x \in \overline{S}, \forall y \in S$, the sets $L_1(x,r)$ and $L_2(y,r)$ are bounded where

$$\begin{aligned} L_1(x,r) &= \{y \in S / D_h(x,y) \leq r\}, \\ L_2(y,r) &= \{x \in \overline{S} / D_h(x,y) \leq r\}. \end{aligned} \tag{10}$$

(i) H_4: if $\{y^k\}_k \subset S$ is such that $y^k \longrightarrow y^* \in \overline{S}$, then,

$$D_h(y^*, y^k) \longrightarrow 0. \tag{11}$$

H_5: if $\{x^k\}_k$ and $\{y^k\}_k$ are two sequences of S such that $D_h(x^k, y^k) \longrightarrow 0$ and $x^k \longrightarrow x^* \in S$, then

$$y^k \longrightarrow x^*. \tag{12}$$

Definition 1

(i) h: $\overline{S} \longrightarrow R$ is a Bregman type function on S or "D-function" if h verifies H_1, H_2, H_3, H_4, and H_5.

(ii) $D_h(.,.)$ is called entropic distance if h is a Bregman function.

We put

$A(S) = \{h\colon \overline{S} \longrightarrow R$ verifying H_1 and $H_2\}$

$B(S) = \{h\colon \overline{S} \longrightarrow R$ verifying H_1, H_2, H_3, H_4, and $H_5\}$.

Theorem 1 (see [17]). *Let $f \in \Gamma_0(R^d)$ and $h \in A(S)$ such that $\operatorname{dom} f \cap \overline{S} \neq \phi$.*

If one of the two following conditions is verified,

(i) $\inf_{\overline{S}} f > -\infty$ and h verifies H_3

(ii) $\operatorname{Im}\nabla h = R^d$,

then for all $x \in S$ and for all $\lambda > 0$, the function $u \longrightarrow f(u) + \lambda^{-1} D_h(u,v)$ has a unique minimum point on $\overline{S}$.

Definition 2. f and h verify the hypothesis of Theorem 1.

(i) The entropic approximation of f compared to h, of parameter $\lambda\,(\lambda > 0)$, is the function defined by

$$f_{h\lambda}(x) := \inf_{y \in \overline{S}} \{f(y) + \lambda^{-1} D_h(y,x)\}, \quad \forall x \in S. \tag{13}$$

(ii) The application entropic proximal of f comparing to h, of parameter λ, is the operator defined by

$$h_\lambda^f(x) := \operatorname{prox}_{\lambda f}^h(x) := \arg\min_{y \in \overline{S}} \{f(y) + \lambda^{-1} D_h(y,x)\}, \quad \forall x \in S. \tag{14}$$

Proposition 1 (see [17]). *Let $h \in A(S)$ and $f \in \Gamma_0(R^d)$ such that*

(a) $ri\ (dom\ f) \cap S \neq \phi$

(b) $\operatorname{Im}\nabla h = R^d$

Then, $\forall x \in S, \forall \lambda > 0$.

$$h_\lambda^f(x) \in S, \tag{15}$$

$$\inf_S f_{h\lambda} = \inf_S f, \tag{16}$$

$$\frac{\nabla h(x) - \nabla h\left(h_\lambda^f(x)\right)}{\lambda} \in \partial f\left(h_\lambda^f(x)\right), \tag{17}$$

$$f_{h\lambda}(x) \leq f_{h\mu}(x) \leq f(x), \quad \forall \mu\colon 0 < \mu \leq \lambda. \tag{18}$$

Proposition 2 (see [17]). *We suppose that h and f verify the conditions of Proposition 1.*

If $\inf(f) > -\infty$ and h verify H_3, then $h_\lambda^f\colon S \longrightarrow S$ is a continuous application.

Proposition 3 (see [17]). *We suppose that h and f verify the hypothesis of Proposition 2.*

If h is twice continuously differentiable on S and $D_h(.,.)$ and jointly convex, then $f_{h\lambda}$ is continually differentiable and convex such that $\forall x \in S$:

$$\nabla f_{h\lambda}(x) = \lambda^{-1} H(x)\left(x - h_\lambda^f(x)\right), \tag{19}$$

where $H = \nabla^2 h$.

Proposition 4. *We suppose that h and f verify the hypothesis of the Proposition 3. If H is defined positive, then*

$$\operatorname{Arg}\min_S f = \operatorname{Arg}\min_S f_{h\lambda}. \tag{20}$$

Proof. Let $u^* \in Arg\min_S f_{h\lambda}$.

$$\begin{aligned} f_{h\lambda}(u^*) = \inf_S f_{h\lambda} &\Longleftrightarrow 0 \in \partial f_{h\lambda}(u^*) \\ &\Longleftrightarrow 0 = \nabla f_{h\lambda}(u^*) \\ &\Longleftrightarrow H(u^*)\left(u^* - h_\lambda^f(u^*)\right). \end{aligned} \tag{21}$$

Since H is defined positive, we deduct then that $u^* = h_\lambda^f(u^*)$. From (17), we have

$$u^* = h_\lambda^f(u^*) \Longrightarrow 0 \in \partial f(u^*) \Longrightarrow u^* \in \arg\min_S f. \tag{22}$$

We get then $\operatorname{Arg}\min_S f_{h\lambda} \subset \operatorname{Arg}\min_S f$.

Reciprocally, let x^* such that $f(x^*) = \inf_S f$. From (16) and (18), we have

$$f(x^*) = \inf_S f_{h\lambda} \le f_{h\lambda}(x^*) \le f(x^*); \quad (23)$$

thus, we have $f(x^*) = \inf_S f_{h\lambda} = f_{h\lambda}(x^*)$, which completes the demonstration.

Some examples of Bregman functions are given below. □

Example 1. If $S_0 = R^d$ and $h_0(x) = (1/2)\|x\|^2$, then

$$D_{h_0}(x, y) = \frac{1}{2}\|x - y\|^2. \quad (24)$$

Example 2. If $S_1 = R^d_{++} := \{x \in R^d / x_i > 0, i = 1, \dots, d\}$ and

$$h_1(x) = \sum_{i=1}^{i=d} x_i \log x_i - x_i; \quad \forall x \in \overline{S}_1, \quad (25)$$

with the convention $0 \log 0 = 0$, then

$$D_{h_1}(x, y) = \sum_{i=1}^{d} x_i \log \frac{x_i}{y_i} + y_i - x_i, \quad \forall (x, y) \in \overline{S}_1 X S_1. \quad (26)$$

Example 3. If $S_2 = [-1, 1]^d$ and $h_2(x) = -\sum_{i=1}^{i=d} \sqrt{1 - x_i^2}$, then

$$D_{h_2}(x, y) = h_2(x) + \sum_{i=1}^{d} \frac{1 - x_i y_i}{\sqrt{1 - y_i^2}}, \quad \forall (x, y) \in \overline{S_2} X S_2. \quad (27)$$

We easily verify that $h_i \in B(S_i), i = 0, 1, 2$.

3. Analysis of the Diagonal Bregman Proximal Method

In this paragraph, we assume the following:

(A): $h \in B(S)$: $\operatorname{Im}\nabla h = R^d$ and $\overline{\operatorname{domf}} \subset S$

(B): $f, f^k \in \Gamma_0(R^d)$: $\overline{\operatorname{domf}^k} \subset S$, $k = 1, 2, \dots$

(C): $\lim\inf (\inf f^k) > -\infty$

From (15), we can then construct the sequence $\{x^k\}_k$ defined by (Algorithm 1):

In what follows, we will derive a convergence result (Theorem 2) for the DPMD framework. First, we need to establish a few technical results.

Lemma 1 (see [20]). *Let f_1, f_2 be two functions of $\Gamma_0(R^d)$ if there exists $\overline{x} \in domf_1$ in which f_2 is finite and continuous, then for $\varepsilon > 0$, for all $y \in domf_1 \cap domf_2$,*

$$\partial_\varepsilon (f_1 + f_2)(y) = \bigcup_{\varepsilon_1 + \varepsilon_2 = \varepsilon, \varepsilon_1 \ge 0, \varepsilon_2 \ge 0} \partial_{\varepsilon_1} f_1(y) + \partial_{\varepsilon_2} f_2(y). \quad (28)$$

Definition 3. The sequence $\{(\lambda_k; a_k; b_k; c_k; d_k)\}_k \in R^{+*} X S^4$ verifies the K-property only if the following properties are verified:

K_1 : $\exists \underline{\lambda} > 0, \forall k, \lambda_k \ge \underline{\lambda}$.

K_2 : $\{a_k\}$ is bounded and $\operatorname{Adh}\{a_k\} \subset S$.

K_3 : $D_h(a_k, b_k) \longrightarrow 0$.

K_4 : $D_h(a_k, c_k) \longrightarrow 0$.

K_5 : $d_k = (\nabla h(b_k) - \nabla h(c_k))/\lambda_k$.

Lemma 2. *If the sequence $\{(\lambda_k; a_k; b_k; c_k; d_k)\}_k$ verifies the K-property, then $d_k \longrightarrow 0$.*

Proof. If the sequence $\{d_k\}$ does not tend to zero, then it exists that $M > 0$ and the subsequence $\{d_{k_i}\}$ of $\{d_k\}$ such that

$$\forall k_i, \left\| d_{k_i} \right\| > M. \quad (29)$$

The sequence $\{a_{k_i}\}$ is bounded and Adh $\{a_{k_i}\} \subset S$; it exists that the subsequence $\{a_{k_j}\}$ of $\{a_{k_i}\}$ and $u^* \in S$ such that $a_{k_j} \longrightarrow u^*$. $D_h(a_{k_j}, b_{k_j}) \longrightarrow 0$ and $D_h(a_{k_j}, c_{k_j}) \longrightarrow 0$ allow to write, from H_5, $b_{k_j} \longrightarrow u^*$ and $c_{k_j} \longrightarrow u^*$. On the other hand,

$$0 \le \left\| d_{k_j} \right\| = \left\| \frac{\nabla h\left(b_{k_j}\right) - \nabla h\left(c_{k_j}\right)}{\lambda_{k_j}} \right\| \le \underline{\lambda}^{-1} \left\| \nabla h\left(b_{k_j}\right) - \nabla h\left(c_{k_j}\right) \right\|, \quad (30)$$

∇h is continuous on S, then $\nabla h(b_{k_j}) - \nabla h(c_{k_j}) \longrightarrow 0$. It follows that $\|d_{k_j}\| \longrightarrow 0$. $\{d_{k_j}\}$ is a subsequence of $\{d_{k_i}\}$, from with the entropic proximal method (29), we have $0 \ge M > 0$, so $d_k \longrightarrow 0$.

Lets consider now the function $h_{u,\lambda}$ defined by $h_{u,\lambda}$: $\overline{S} \longrightarrow R$, $\forall \lambda > 0, \forall u \in S$.

$$h_{u,\lambda}(x) = \lambda^{-1} D_h(x, u), \quad \forall x \in \overline{S}. \quad (31)$$

□

Proposition 5. $\forall \varepsilon > 0, \forall \lambda > 0, \forall u \in S, \forall x^* \in \overline{S}$.

$$\partial_\varepsilon h_{u,\lambda}(x^*) = \left\{ \frac{z}{z} = \frac{\nabla h(\overline{x}) - \nabla h(u)}{\lambda} \text{ with } \overline{x} \in S \text{ and } D_h(x^*, \overline{x}) \le \lambda\varepsilon \right\}. \quad (32)$$

Proof. $z \in \partial_\varepsilon h_{u,\lambda}(x^*)$.

$$\begin{aligned}
&\Longleftrightarrow h_{u,\lambda}(x) - h_{u,\lambda}(x^*) \ge \langle z, x - x^* \rangle - \varepsilon, \quad \forall x \in \overline{S}, \\
&\Longleftrightarrow \lambda^{-1} [h(x) - h(u) - \langle x - u, \nabla h(u) \rangle - h(x^*) + h(u) \\
&\quad + \langle x^* - u, \nabla h(u) \rangle] \ge \langle z, x - x^* \rangle - \varepsilon, \quad \forall x \in \overline{S}, \\
&\Longleftrightarrow h(x^*) - h(x) - \langle x^* - x, \nabla h(u) \rangle \\
&\quad \le \langle \lambda z, x^* - x \rangle + \lambda\varepsilon, \quad \forall x \in \overline{S},
\end{aligned} \quad (33)$$

which is equivalent to

$$h(x^*) - h(x) - \langle x^* - x, \nabla h(u) + \lambda z \rangle \le \lambda\varepsilon. \quad (34)$$

According to (A), it exists that $\overline{x} \in S$ such that

(1) Input: $x^0 \in S$
(2) Choose $\overline{\lambda} \geq \lambda_k \geq \underline{\lambda} > 0$ and $\varepsilon_k \geq 0$, and find $x^k \in S$, such that $x^k \in \varepsilon_k - \operatorname{Arg\,min}\{f^k(\cdot) + \lambda_k^{-1} D_h(\cdot, x^{k-1})\}$.
(3) Set $k \longleftarrow k+1$ and go to step 2

ALGORITHM 1: DPMD.

$$\nabla h(u) + \lambda z = \nabla h(\overline{x}), \tag{35}$$

which means

$$z = \frac{\nabla h(\overline{x}) - \nabla h(u)}{\lambda} \quad \exists \overline{x} \in S. \tag{36}$$

Replacing in (34) x by $\overline{x}$, we get

$$D_h(x^*, \overline{x}) \leq \lambda\varepsilon. \tag{37}$$

Finally

$$\partial_\varepsilon h_{u,\lambda}(x^*) \subset \left\{ \frac{z}{z} = \frac{\nabla h(\overline{x}) - \nabla h(u)}{\lambda} \text{ with } \overline{x} \in S \text{ and } D_h(x^*, \overline{x}) \leq \lambda\varepsilon \right\}. \tag{38}$$

Conversely, let z such as

$$z = \frac{\nabla h(\overline{x}) - \nabla h(u)}{\lambda},$$

$$D_h(x^*, \overline{x}) \leq \lambda\varepsilon,$$

$$\begin{aligned} D_h(x^*, \overline{x}) \leq \lambda\varepsilon &\Longrightarrow h(x^*) - h(\overline{x}) - \langle x^* - \overline{x}, \nabla h(\overline{x})\rangle \leq \lambda\varepsilon \leq \lambda\varepsilon \\ &\quad + D_h(x, \overline{x}), \\ &\Longrightarrow h(x^*) - h(\overline{x}) - \langle x^* - \overline{x}, \nabla h(\overline{x})\rangle - h(x) \\ &\quad + h(\overline{x}) + \langle x - \overline{x}, \nabla h(\overline{x})\rangle \leq \lambda\varepsilon, \\ &\Longrightarrow h(x^*) - h(x) - \langle x^* - x, \nabla h(\overline{x})\rangle \leq \lambda\varepsilon. \end{aligned} \tag{39}$$

Replacing $\nabla h(\overline{x})$ by $\nabla h(u) + \lambda z$, we get (34). According to what precedes,

$$(34) \Longleftrightarrow z \in \partial_\varepsilon h_{u,\lambda}(x^*), \tag{40}$$

which establishes the desired equality. □

Definition 4.

$$\begin{gathered} \forall \lambda > 0, \forall \rho \geq 0, \forall f, g \in \Gamma_0(R^d), \\ d^\rho_{h,\lambda}(f, g) := \sup_{\|x\| \leq \rho, x \in S} \left| f_{h\lambda}(x) - g_{h\lambda}(x) \right|. \end{gathered} \tag{41}$$

Theorem 2. *We assume that*

(i) $\sum_k \varepsilon_k + 2d_{k,\rho} < +\infty, \forall \rho \geq 0$, *where*

$$d_{k,\rho} := d^\rho_{h,\underline{\lambda}}(f^k, f). \tag{42}$$

(ii) The sequence $\{x^k\}$ *generated by DPMD is bounded. Then*

(a) $f^k(x^k) \longrightarrow \inf f$

(b) Moreover, if f *and* h *verify the conditions of Proposition 4, then*

$$\operatorname{Adh}\{x^k\} \subset \operatorname{Arg\,min} f. \tag{43}$$

Proof.

$$\begin{aligned} &x^k \in \varepsilon_k - \operatorname{Arg\,min}\{f^k(u) + \lambda_k^{-1} D_h(u, x^{k-1})\} \\ &\quad \Longrightarrow f^k(x^k) + \lambda_k^{-1} D_h(x^k, x^{k-1}) \leq f^k_{h\lambda_k}(x^{k-1}) + \varepsilon_k, \end{aligned} \tag{44}$$

according to (18), we can write

$$f^k_{h\underline{\lambda}}(x^k) + \lambda_k^{-1} D_h(x^k, x^{k-1}) \leq f^k_{h\underline{\lambda}}(x^{k-1}) + \varepsilon_k. \tag{45}$$

The sequence $\{x^k\}$ is bounded; let $\rho \geq 0$ such that

$$\forall k, \left\|x^k\right\| \leq \rho. \tag{46}$$

Considering (45),

$$\lambda_k^{-1} D_h(x^k, x^{k-1}) + f_{h\underline{\lambda}}(x^k) \leq f_{h\underline{\lambda}}(x^{k-1}) + \varepsilon_k + 2d_{k,\rho}. \tag{47}$$

Therefore,

$$f_{h\underline{\lambda}}(x^k) \leq f_{h\underline{\lambda}}(x^{k-1}) + \varepsilon_k + 2d_{k,\rho}. \tag{48}$$

So, from (i), we have

$$\lim f_{h\underline{\lambda}}(x^k) = l \geq \inf f_{h\underline{\lambda}} = \inf f \geq -\infty. \tag{49}$$

On one hand,

$$\begin{gathered} f^k(x^k) \leq f^k_{h\underline{\lambda}}(x^{k-1}) + \varepsilon_k \\ \Downarrow \\ f^k(x^k) \leq f_{h\underline{\lambda}}(x^{k-1}) + d_{k,\rho} + \varepsilon_k, \end{gathered} \tag{50}$$

on the other hand, we have

$$f_{h\underline{\lambda}}(x^k) - d_{k,\rho} \leq f^k(x^k); \tag{51}$$

finally, the two previous inequalities make it possible to write

$$f_{h\underline{\lambda}}(x^k) - d_{k,\rho} \leq f^k(x^k) \leq f_{h\underline{\lambda}}(x^{k-1}) + d_{k,\rho} + \varepsilon_k. \tag{52}$$

If $l = -\infty$, then $\inf f = -\infty$ and $f^k(x^k) \longrightarrow -\infty$. So,

$$f^k(x^k) \longrightarrow \inf f. \tag{53}$$

If $l > -\infty$, then, from (52),

$$\lim f^k(x^k) = \lim f_{h\underline{\lambda}}(x^k) = l. \tag{54}$$

Let us show that $l = \inf f$, from (45), $\lambda_k^{-1} D_h(x^k, x^{k-1}) \longrightarrow 0$ when $k \longrightarrow +\infty$. As $\underline{\lambda} \leq \lambda_k \leq \overline{\lambda}$, we have

$$D_h\left(x^k, x^{k-1}\right) \longrightarrow 0. \tag{55}$$

On the other hand,

$$\begin{aligned} x^k &\in \varepsilon_k - \operatorname{Arg\,min}\left\{f^k(u) + \lambda_k^{-1} D_h\left(u, x^{k-1}\right)\right\} \\ &\Longleftrightarrow 0 \in \partial_{\varepsilon_k}\left[f^k(\cdot) + \lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right)\right]\left(x^k\right). \end{aligned} \tag{56}$$

From Lemma 1, there exists $\varepsilon_{k_1}, \varepsilon_{k_2} \geq 0$ such that $\varepsilon_{k_1} + \varepsilon_{k_2} = \varepsilon_k$ and

$$0 \in \partial_{\varepsilon_{k_1}} f^k\left(x^k\right) + \partial_{\varepsilon_{k_2}}\left(\lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right)\right)\left(x^k\right). \tag{57}$$

Since $\partial_\varepsilon f^k$ increases with ε, we have

$$0 \in \partial_{\varepsilon_k} f^k\left(x^k\right) + \partial_{\varepsilon_k}\left(\lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right)\right)\left(x^k\right). \tag{58}$$

Therefore, there exists $z_k \in \partial_{\varepsilon_k} f^k(x^k)$ such that

$$-z_k \in \partial_{\varepsilon_k}\left(\lambda_k^{-1} D_h\left(\cdot, x^{k-1}\right)\right)\left(x^k\right). \tag{59}$$

From Proposition 5, there exits $\overline{x}^k \in S$ such that

$$\begin{aligned} -z_k &= \frac{\nabla h\left(\overline{x}^k\right) - \nabla h\left(x^{k-1}\right)}{\lambda_k}, \\ D_h\left(x^k, \overline{x}^k\right) &\leq \lambda_k \varepsilon_k. \end{aligned} \tag{60}$$

Finally, there exists $\left\{\overline{x}^k\right\}$ such that

$$z_k = \frac{\nabla h\left(x^{k-1}\right) - \nabla h\left(\overline{x}^k\right)}{\lambda_k} \in \partial_{\varepsilon_k} f^k\left(x^k\right) \text{with}\, D_h\left(x^k, \overline{x}^k\right) \leq \lambda_k \varepsilon_k. \tag{61}$$

From (55) and (61), we have

$$\begin{cases} z_k = \dfrac{\nabla h\left(x^{k-1}\right) - \nabla h\left(\overline{x}^k\right)}{\lambda_k} \in \partial_{\varepsilon_k} f^k\left(x^k\right), \\ D_h\left(x^k, \overline{x}^k\right) \longrightarrow 0, \\ D_h\left(x^k, x^{k-1}\right) \longrightarrow 0. \end{cases} \tag{62}$$

Since $\operatorname{Adh}\ \{x^k\} \subset \cup \overline{\operatorname{dom} f^k} \subset S$, the sequence $\left\{(\lambda_k; x^k; \overline{x}^k; x^{k-1}; z_k)\right\}_k$ verifies then the K-property. From Lemma 2, $z_k \longrightarrow 0$. On the other hand, for all $y \in S$,

$$\begin{aligned} f^k(y) \geq f^k\left(x^k\right) + \langle z_k, y - x^k\rangle - \varepsilon_k &\Longrightarrow f^k(y) + \underline{\lambda}^{-1} D_h(y, x) \\ &\geq f^k\left(x^k\right) + \langle z_k, y - x^k\rangle - \varepsilon_k, \end{aligned} \tag{63}$$

there exists $\widehat{x}^k$ such that for all $k \in N$,

$$\begin{aligned} \inf\left\{f^k(y) + \underline{\lambda}^{-1} D_h(y, x)\right\} &= f^k_{h\underline{\lambda}}(x) \\ &= f^k\left(\widehat{x}^k\right) + \underline{\lambda}^{-1} D_h\left(\widehat{x}^k, x\right). \end{aligned} \tag{64}$$

By replacing y by $\widehat{x}^k$ in (63), we get

$$f^k\left(\widehat{x}^k\right) + \underline{\lambda}^{-1} D_h\left(\widehat{x}^k, x\right) \geq f^k\left(x^k\right) + \langle z_k, \widehat{x}^k - x^k\rangle - \varepsilon_k. \tag{65}$$

It is still

$$f^k_{h\underline{\lambda}}(x) \geq f^k\left(x^k\right) + \langle z_k, \widehat{x}^k - x^k\rangle - \varepsilon_k. \tag{66}$$

$\left\{\widehat{x}^k\right\}$ is bounded. Indeed,

$$-\infty < \inf_k \inf_S f^k \leq f^k\left(\widehat{x}^k\right), \tag{67}$$

so it exists $K_1 \in R$ such that $K_1 \leq f^k(\widehat{x}^k)$.

From (i),

$$d_{k,\rho} \longrightarrow 0 \Longrightarrow \left\{f^k_{h\underline{\lambda}}(x)\right\} \text{is convergent}, \tag{68}$$

so

$$\begin{aligned} \exists K_2 \in R&: K_2 \geq K_1, \\ f^k_{h\underline{\lambda}}(x) &\leq K_2. \end{aligned} \tag{69}$$

From (64), we have

$$D_h\left(\widehat{x}^k, x\right) \leq \underline{\lambda}\left(K_2 - K_1\right). \tag{70}$$

From H_3, $\left\{\widehat{x}^k\right\}$ is bounded. Going to the limit in (66), we have

$$f_{h\underline{\lambda}}(x) \geq \lim f^k\left(x^k\right) = l, \tag{71}$$

then,

$$\begin{aligned} f(x) \geq l, \quad \forall x &\Longrightarrow \inf f \geq l, \\ (52) &\Longrightarrow f_{h\underline{\lambda}}\left(x^k\right) \leq f^k\left(x^k\right) + d_{k,\rho}, \\ &\Longrightarrow \inf f = \inf f_{h\underline{\lambda}} \leq f^k\left(x^k\right) + d_{k,\rho}, \\ &\Longrightarrow \inf f \leq \lim f^k\left(x^k\right) = l. \end{aligned} \tag{72}$$

Finally, we have

$$\lim f^k\left(x^k\right) = \inf f. \tag{73}$$

(b) Let $x^* \in \operatorname{Adh}\{x^k\}$, there exists then the subsequence $\{x^{k_i}\}$ of $\{x^k\}$ such that $x^{k_i} \longrightarrow x^*$, we have

$$\begin{aligned} \inf f_{h\underline{\lambda}} &\leq f_{h\underline{\lambda}}(x^*) \leq \lim f_{h\underline{\lambda}}\left(x^{k_i}\right) \leq \lim f^{k_i}\left(x^{k_i}\right) \\ &= \inf f = \inf f_{h\underline{\lambda}} \Longrightarrow \inf f_{h\underline{\lambda}} = f_{h\underline{\lambda}}(x^*). \end{aligned} \tag{74}$$

From (20), we have

$$x^* \in \operatorname{Arg\,min} f_{h\underline{\lambda}} = \operatorname{Arg\,min} f. \tag{75}$$

□

4. Exterior Penalty Coupled with Bregman Proximal Method

Let f_i: $R^d \longrightarrow R, i = 1, \ldots, m$, be the convex function and let C the set of constraints given by

$$C = \left\{x \in R^d \colon f_i(x) \leq 0, i = 1, \ldots, m\right\}. \tag{76}$$

We suppose that C verifies the condition of Slater:

$$\exists x^* \in R^d \colon f_i(x^*) < 0, \quad i = 1, \dots, m. \tag{77}$$

Let us consider the functions of the linear penalty defined by

$$\varphi_1^n(x) = r_n \sum_{i=1}^{m} f_i^+(x), \quad \forall x \in R^d, \forall n \in N^*, \tag{78}$$

and the quadratic exterior penalty defined by

$$\varphi_2^n(x) = r_n \sum_{i=1}^{m} [f_i^+(x)]^2, \quad \forall x \in R^d, \forall n \in N^*, \tag{79}$$

where $a^+ = \max\{0, a\}$ and $\{r_n\}_n$ is an increasing sequence of strictly positive real numbers which tends to $+\infty$.

Let us put $\forall n \in N^*$:

$$\begin{aligned} f &= f_0 + \Psi_C, \\ f_1^n &= f_0 + \varphi_1^n, \\ f_2^n &= f_0 + \varphi_2^n. \end{aligned} \tag{80}$$

In what follows, we assume

(A'): $h \in B(R^d)$: $\mathrm{Im}\nabla h = R^d$

(B'): $\inf_{R^d} f > -\infty$,

so conditions (A), (B), and (C) of Section 3 are verified for f and f_j^n, $j = 1, 2$; $n \in N$.

We give below an estimate of $d_{\lambda,\rho}^h(f_j^n, f)$, $j = 1, 2$.

Proposition 6

(a) $\forall \underline{\lambda} > 0, \forall \rho \geq 0, \exists r_h \geq 0,$

$$d_{\lambda,\rho}^h(f_1^n, f) = 0, \quad \forall n \colon r_n \geq r_h, \forall \lambda \geq \underline{\lambda}. \tag{81}$$

(b) $\forall \underline{\lambda} > 0, \forall \rho \geq 0, \exists \mu_h \geq 0,$

$$d_{\lambda,\rho}^h(f_2^n, f) \leq \frac{\mu_h}{r_n}, \quad \forall n \geq 1, \forall \lambda \geq \underline{\lambda}. \tag{82}$$

Proof. Let $\lambda > 0$ and $x \in R^d$.

$$f_{h\lambda}(x) = \inf\{f + \lambda^{-1} D_h(\cdot, x)\} = \inf_C\{f_0 + \lambda^{-1} D_h(\cdot, x)\}. \tag{83}$$

Since Slater's condition is verified, there exists from Ekeland-Temam [21] (chap 3, Theorem 5.2) multiplicators of Lagrange

$$p_j(\lambda, x) \geq 0; \quad j = 1, \dots, m, \tag{84}$$

such that $\forall y \in R^d$.

$$f_{h\lambda}(x) \leq f_0(y) + \lambda^{-1} D_h(y, x) + \sum_{j=1}^{m} p_j(\lambda, x) f_j(y). \tag{85}$$

From (18), we have $f_{h\lambda}(x) \geq \inf_C f_0$, by replacing y with x^*, we obtain

$$\inf_C f_0 \leq f_0(x^*) + \lambda^{-1} D_h(x^*, x) + \sum_{j=1}^{m} p_j(\lambda, x) f_j(x^*), \tag{86}$$

where x^* verifies (3). On the other hand

$$\sum_{j=1}^{m} p_j(\lambda, x) f_j(x^*) \leq \left[\sum_{j=1}^{m} p_j(\lambda, x)\right] \sup_j f_j(x^*). \tag{87}$$

Let us put

$$\sup_j f_j(x^*) = -c, \quad \text{where } c > 0. \tag{88}$$

It follows that

$$\sum_{j=1}^{m} p_j(\lambda, x) f_j(x^*) \leq -c \|p(\lambda, x)\|_1, \tag{89}$$

where $p(\lambda, x) = (p_1(\lambda, x), \dots, p_m(\lambda, x))$. Therefore,

$$\begin{aligned} D_h(x^*, x) &\leq D_h(x^*, x) + D_h(x, x^*) \\ &= \langle x^* - x, \nabla h(x^*) - \nabla h(x) \rangle \\ &\leq \|x^* - x\| \cdot \|\nabla h(x) - \nabla h(x^*)\|, \end{aligned} \tag{90}$$

which leads to

$$D_h(x^*, x) \leq (\|x\|^* + \|x\|)(\|\nabla h(x^*)\| + \|\nabla h(x)\|). \tag{91}$$

From (86), (89), and (91), we obtain

$$\begin{aligned} \|p(\lambda, x)\|_1 \leq \frac{1}{c} \Big[& f_0(x^*) + \underline{\lambda}^{-1}(\|x^*\| + \|x\|)(\|\nabla h(x^*)\| \\ & + \|\nabla h(x)\|) - \inf_C f_0 \Big]. \end{aligned} \tag{92}$$

(a) From (85),

$$f_{h\lambda}(x) \leq f_0(y) + \underline{\lambda}^{-1} D_h(y, x) + \|p(\lambda, x)\|_1 \sum_{i=1}^{m} f_i^+(y). \tag{93}$$

For $x \in B_\rho$, from (92), we have

$$\begin{aligned} \|p(\lambda, x)\|_1 \leq \frac{1}{c} \Big[& f_0(x^*) + \underline{\lambda}^{-1}(\|x^*\| + \rho) \\ & \cdot \left(\|\nabla h(x^*)\| + \sup_{x \in B_\rho} \|\nabla h(x)\| \right) - \inf_C f_0 \Big] =: r_h. \end{aligned} \tag{94}$$

Thus, for n such as $r_n \geq r_h$,

$$\begin{gathered} f_{h\lambda}(x) \leq f_1^n(y) + \underline{\lambda}^{-1} D_h(y, x), \quad \forall y \in R^p \\ \Downarrow \\ f_{h\lambda}(x) \leq (f_1^n)_{h\lambda}(x), \quad \forall x \in B_\rho. \end{gathered} \tag{95}$$

Conversely,

$$\begin{aligned}(f_1^n)_{h\lambda}(x) &= \inf_{y\in R^p}\left\{f_0(y) + r_n\sum_{i=1}^{m} f_i^+(y) + \lambda^{-1}D_h(y,x)\right\}\\ &\le \inf_{y\in C}\left\{f_0(y) + \lambda^{-1}D_h(y,x)\right\}\\ &= \inf\left\{f(y) + \lambda^{-1}D_h(y,x)\right\}.\end{aligned} \tag{96}$$

Therefore,

$$(f_1^n)_{h\lambda}(x) \le f_{h\lambda}(x). \tag{97}$$

(b) If $\|\cdot\|_m$ indicates the Euclidean norm on R^d, so

$$\sum_{j=1}^{m} p_j(\lambda,x) f_j(x) \le \frac{1}{2r_n}\|p(\lambda,x)\|_m^2 + \frac{r_n}{2}\sum_{i=1}^{m}\left[f_i^+(y)\right]^2; \tag{98}$$

by proceeding as below, we deduce the result.

Let us consider now the methods of the coupled exterior penalties with the entropic proximal method (Algorithm 2): □

Theorem 3. *Let us suppose*

(i) $\sum_{n=1}^{\infty}\varepsilon_n < +\infty$, *and* $\sum_{n=1}^{\infty}(1/r_n) < +\infty$

(ii) $\exists i_0 \ge 1$: f_{i_0} *is coercive*

Then the sequence $\{x_j^n\}_n$ *generated by* $(DBPM)_j$ $j = 1, 2$ *is bounded,* $Adh\{x_j^n\} \subset Arg\min f$, *and*

$$f_j^n\{x_j^n\} \longrightarrow \inf f, \quad j = 1, 2 \ldots. \tag{99}$$

Proof. Let us show that $\{x_1^n\}_n$ is bounded.

$$f_1^n(x_1^n) + \lambda_n^{-1}D_h\left(x_1^n, x_1^{n-1}\right) \le f_1^n(u) + \lambda_n^{-1}D_h\left(u, x_1^{n-1}\right) + \varepsilon_n. \tag{100}$$

By replacing u by x_1^{n-1} in (100), we obtain

$$f_0(x_1^n) + r_n\sum_{i=1}^{m} f_i^+(x_1^n) \le f_0\left(x_1^{n-1}\right) + r_n\sum_{i=1}^{m} f_i^+\left(x_1^{n-1}\right) + \varepsilon_n. \tag{101}$$

Let $R = \inf_{R^d} f_0$, we have

$$\frac{f_0(x_1^n) - R}{r_n} + \sum_{i=1}^{m} f_i^+(x_1^n) \le \frac{f_0(x_1^{n-1}) - R}{r_n} + \sum_{i=1}^{m} f_i^+\left(x_1^{n-1}\right) + \frac{\varepsilon_n}{r_n}, \tag{102}$$

$\{r_n\}$ is increasing, so we put

$$A_n = \frac{f_0(x_1^n) - R}{r_n} + \sum_{i=1}^{m} f_i^+(x_1^n), \tag{103}$$

we deduce

$$A_n \le A_{n-1} + \frac{\varepsilon_n}{r_n}. \tag{104}$$

which leads to

$$A_n \le A_0 + \sum_{k=1}^{n}\frac{\varepsilon_k}{r_k}. \tag{105}$$

Let ε such as $\varepsilon_k \le \varepsilon$, since $(f_0(x_1^n) - R/r_n) > 0$, we have

$$f_{i_0}(x_1^n) \le A_0 + \sum_{k=1}^{\infty}\frac{\varepsilon}{r_k}. \tag{106}$$

From (i) and (ii), we deduce that the sequence $\{x_1^n\}_n$ is bounded, so by application of Theorem 2 and Proposition 6, the result is immediate.

In a similar way, we deduce the result for $j = 2$. □

5. Example

Let us consider the following optimization problem:

$$(P_1): \begin{cases}\min\langle a, x\rangle,\\ \|x\|^2 \le b,\end{cases} \tag{107}$$

where $a \in R^d$ and $b \in R^{+*}$.

The $(DBPM)_2$ algorithm can be applied to solve (P_1). We take

$$\begin{aligned}&\forall n, \varepsilon_n = 0,\\ &r_n = 2^n,\\ &\bar{\lambda} \ge \lambda_n \ge \underline{\lambda} > 0.\end{aligned} \tag{108}$$

Let us consider the function h: $R^d \longrightarrow R$ defined by

$$h(x) = \frac{1}{p}\sum_{i=1}^{i=d}|x_i|^p; \quad p \ge 2; \forall x \in R^d. \tag{109}$$

We easily show that $h \in B(R^d)$ and which checks (A') Let us put $\forall x \in R^d$:

$$\begin{aligned}&f_0(x) = \langle a, x\rangle,\\ &f_1(x) = \left\|x^2\right\| - b,\\ &\varphi^n(x) = \varphi_2^n(x) = r_n\left[f_1^+(x)\right]^2, \quad \forall n \in N^*.\end{aligned} \tag{110}$$

$C = \{x \in R^d: f_1(x) \le 0\}$ is compact and f_0 is continuous, so

$$\begin{aligned}&\operatorname{Arg}\min f \ne \emptyset,\\ &\inf_{R^d} f = \inf_{C} f_0 > -\infty.\end{aligned} \tag{111}$$

We have

$$\begin{aligned}f &= f_0 + \Psi_C,\\ f^n &:= f_2^n = f_0 + \varphi^n.\end{aligned} \tag{112}$$

The sequence $\{x^n\}$ generated by the $(DBPM)_2$ algorithm is defined by $x^0 \in R^d$ and

$$x^n \in \operatorname{Arg}\min\left\{f^n(\cdot) + \lambda_n^{-1}D_h\left(\cdot, x^{n-1}\right)\right\}, \quad n \ge 1. \tag{113}$$

By writing the condition of optimality, we have

(1) Input: $x_j^0 \in R^d$
(2) Choose $\overline{\lambda} \geq \lambda_n \geq \underline{\lambda} > 0$ and $\varepsilon_n \geq 0$, and find $x_j^n \in S$, such that $x_j^n \in \varepsilon_n - \operatorname{Arg\,min}\left\{f_j^n(\cdot) + \lambda_n^{-1} D_h(\cdot, x_j^{n-1})\right\}$.
(3) Set $n \longleftarrow n+1$ and go to step 2

Algorithm 2: $(\text{DBPM})_j$, $j = 1,2$.

$$\nabla f^n(x^n) + \lambda_n^{-1}\left(\nabla h(x^n) - \nabla h\left(x^{n-1}\right)\right) = 0. \tag{114}$$

On the other hand,

$$\nabla\left[f_1^+\right]^2(x) = 2f_1^+(x)\nabla f_1(x) = 4f_1^+(x)x. \tag{115}$$

Then,

$$2^{n+2} f_1^+(x^n)x^n + \lambda_n^{-1}\nabla h(x^n) = \lambda_n^{-1}\nabla h\left(x^{n-1}\right) - a, \tag{116}$$

where

$$(\nabla h(x^n))_i = \operatorname{sign}(x_i^n)\left|x_i^n\right|^{p-1}, \quad i = 1, \ldots, d, \tag{117}$$

f_1 is coercive, so by applying Theorem 3, we have

$$\begin{aligned} \operatorname{Adh}\{x^n\} \subset \operatorname{Arg\,min} f &= \left\{x \in R^d\colon f_0(x) = \inf_C f_0\right\}, \\ f^n\{x^n\} &\longrightarrow \inf_C f_0. \end{aligned} \tag{118}$$

Remarks 1

(i) The convergence performance of the $\{x^n\}$ can be discussed according to the parameter p. We take note that for $p = 2$,

$$h(x) = \frac{1}{2}\|x\|^2. \tag{119}$$

(ii) Let

$$f_0(x) = \frac{1}{2} < Ax, x > - < c, x >, \quad \forall x \in R^d, \tag{120}$$

where A is matrix symmetric definite positive and $c \in R^d$. Previously developed methods can solve optimization problems of the type

$$(P_2)\colon \begin{cases} \min \dfrac{1}{2}\langle Ax, x\rangle - \langle c, x\rangle, \\ \|x\|^2 \leq b. \end{cases} \tag{121}$$

6. Conclusion

The class of the methods studied in this work constitutes a unified framework for several existing methods that solve convex optimization problems with and without constraints while providing others, more precisely.

(i) For $h(\cdot) = (1/2)\|\cdot\|^2$, DBPM coincides with the diagonal proximal method DPM studied by Alart and Lemaire [1].

(ii) If $h(\cdot) = (1/2)\|\cdot\|^2$ in DBPM, $j = 1, 2$, we find then the methods of penalization studied by Auslender [2].

(iii) If $f^k = f \forall k$, DBPM appears as an inexact version of BPM and solves the problem of convex optimization without constraints:

$$(P')\colon \min\left\{f(x), x \in R^d\right\}, \tag{122}$$

the convergence of this version is included in our analysis and responds to the question asked by Eckstein in [15].

(i) If $f^k = f$ and $\varepsilon_k = 0, \forall k$ in DBPM, we find then BPM studied by [15–18].

(ii) If $f^k = f$ and $h(\cdot) = (1/2)\|\cdot\|^2$ in DBPM, we find then PM studied by [6, 8–13].

(iii) If $f^k = f$ and $h_1(x) = \sum_{i=1}^{i=d} x_i \log x_i - x_i$; $\forall x \in \overline{S_1}$, DBPM allows to minimize f on

$$S_1 = R_{++}^d := \left\{x \in R^d / x_i > 0, i = 1, \ldots, d\right\}. \tag{123}$$

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] P. Alart and B. Lemaire, "Penalization in non-classical convex programming via variational convergence," *Mathematical Programming*, vol. 51, no. 1–3, pp. 307–331, 1991.
[2] A. Auslender, J. P. Crouzeix, and P. Fedit, "Penalty-proximal methods in convex programming," *Journal of Optimization Theory and Applications*, vol. 55, no. 1, pp. 1–21, 1987.
[3] P. Fedit, "Contribution aux méthodes numériques en programmation mathématique non différentiable," Thèse de Troisième Cycle, Université de Clermont II, Clermont-Ferrand, France, 1985.
[4] A. A. Kaplan, "On a convex programming method with internal regularization," *Soviet Mathematics Doklady*, vol. 19, no. 4, pp. 795–799, 1978.

[5] S. Lian and J. Han, "Smoothing approximation to the square-order exact penalty functions for constrained optimization," *Journal of Applied Mathematics*, vol. 2013, Article ID 568316, 7 pages, 2013.

[6] B. Lemaire, "About the convergence of the proximal method , proceedings 6th French German conference on optimisation," in *Advances in Optimization Lecture Notes*, pp. 39–51, Springer-Verlag, Berlin, Germany, 1992.

[7] Mouallif and Tossing, "Une méthode de pénalisation exponentielle associée à une régulrisation proximale," *Bulletin de la Société royale des sciences de Liège*, vol. 56, no. 2, pp. 181–192, 1987.

[8] A. Auslender, "Numerical methods for non-differentiable convex optimization," *Mathematical Programming Studies*, vol. 30, pp. 102–126, 1987.

[9] O. Gouler, "Ergodic convergence in proximal point algorithms with Bregman functions in advances in optimization and approximation," in *Nonconvex Optimization and Its Applications*, pp. 155–165, Kluwer Academic, Dordrecht, The Netherlands, 1994.

[10] B. Lemaire, "Coupling optimization methods and variational convergence," in *Trends in Mathematical Optimization International Series of Numerical Mathematics*, pp. 163–179, Birkhauser, Basel, Switzerland, 1988.

[11] B. Lemaire, "The proximal algorithm," *International Series of Numerical Mathematics*, vol. 87, pp. 73–87, 1989.

[12] B. Martinet, "Perturbation des méthodes d'optimisation—application," *RAIRO Analyse Numérique*, vol. 12, no. 2, pp. 153–171, 1976.

[13] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAMJ Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1975.

[14] L. M. Bregman, "A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 3, pp. 200–217, 1967.

[15] J. Eckstein, "Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 202–226, 1993.

[16] A. Hamdi, M. A. Noor, and A. A. Mukheimer, "Convergence of a proximal point algorithm for solving minimization problems," *Journal of Applied Mathematics*, vol. 2012, Article ID 142862, 13 pages, 2012.

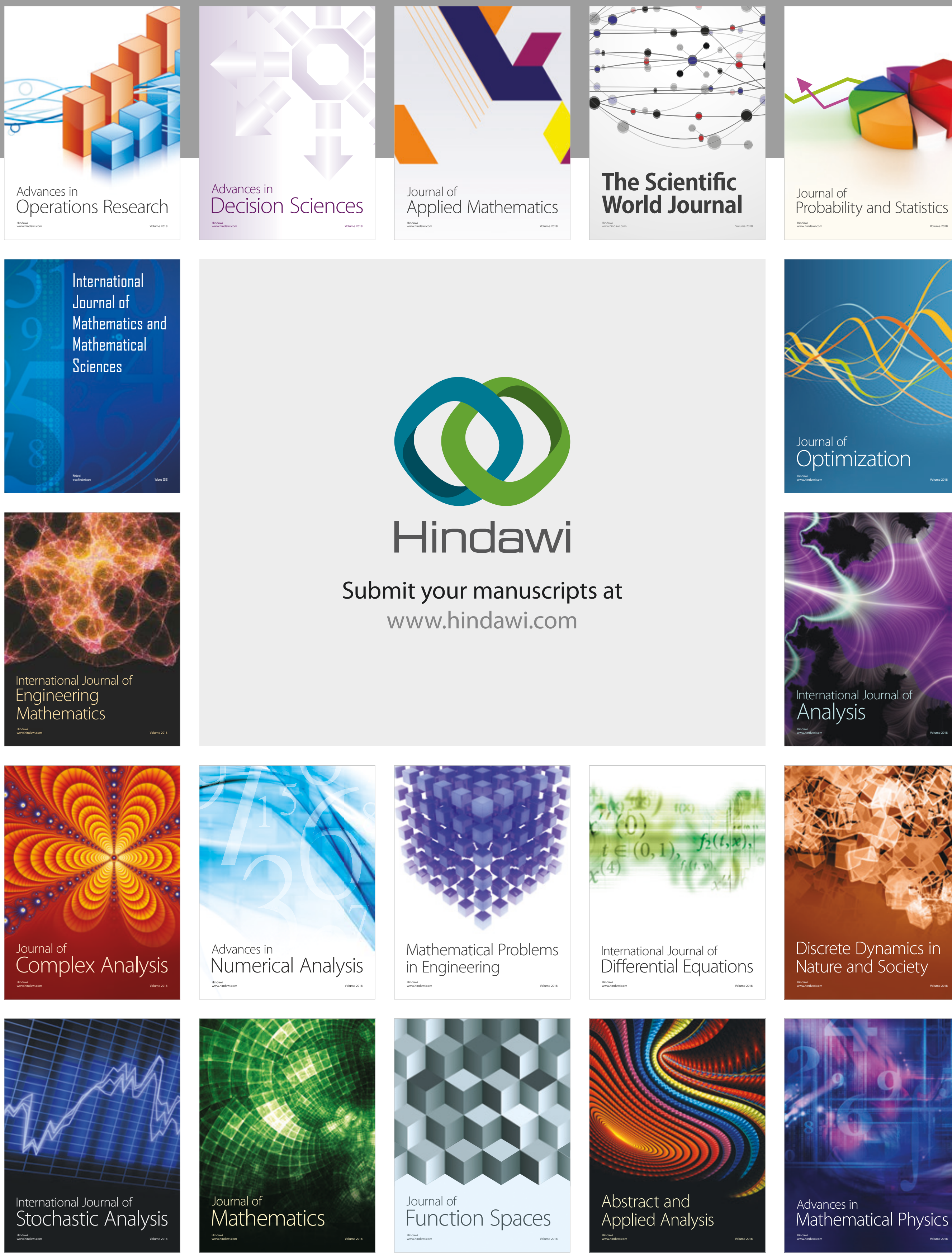
[17] S. Kabbadj, *Méthodes Proximales Entropiques*, Université Montpellier II, Montpellier, France, thesis, 1994.

[18] M. Teboulle and C. Gong, "Convergence analysis of a proximal-like minimization algorithm using Bregman function," *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 538–543, 1993.

[19] J. J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bulletin de la Société mathématique de France*, vol. 79, pp. 273–299, 1965.

[20] J. B. Hiriart-Urruty, "ε-subdifferentiel calculs, convex analysis and optimization," in *Research Notes in Mathematics Series*, Vol. 57, Pitman Publishers, London, UK, 1982.

[21] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels*, Dunod-Gautier-Villars, Paris, France, 1974.


Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
The Scientific World Journal
Journal of Probability and Statistics
International Journal of Mathematics and Mathematical Sciences
Hindawi
Submit your manuscripts at
www.hindawi.com
Journal of Optimization
International Journal of Engineering Mathematics
International Journal of Analysis
Journal of Complex Analysis
Advances in Numerical Analysis
Mathematical Problems in Engineering
International Journal of Differential Equations
Discrete Dynamics in Nature and Society
International Journal of Stochastic Analysis
Journal of Mathematics
Journal of Function Spaces
Abstract and Applied Analysis
Advances in Mathematical Physics